

This is a repository copy of *Cochrane Qualitative and Implementation Methods Group Guidance Paper 3 : Methods for Assessing Evidence on Intervention Implementation*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/126118/>

Version: Accepted Version

Article:

Cargo, M, Harris, J., Pantoja, Tomas et al. (7 more authors) (2017) Cochrane Qualitative and Implementation Methods Group Guidance Paper 3 : Methods for Assessing Evidence on Intervention Implementation. *Journal of Clinical Epidemiology*. ISSN 0895-4356

<https://doi.org/10.1016/j.jclinepi.2017.11.028>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Accepted Manuscript

Cochrane Qualitative and Implementation Methods Group Guidance Paper 3: Methods for Assessing Evidence on Intervention Implementation

Assoc/Prof Margaret Cargo, Dr Janet Harris, Assoc/Prof Tomas Pantoja, Dr Andrew Booth, Prof Angela Harden, Assoc/Prof Karin Hannes, Prof James Thomas, Dr Kate Flemming, Dr Ruth Garside, Prof Jane Noyes

PII: S0895-4356(17)31334-3

DOI: [10.1016/j.jclinepi.2017.11.028](https://doi.org/10.1016/j.jclinepi.2017.11.028)

Reference: JCE 9546

To appear in: *Journal of Clinical Epidemiology*

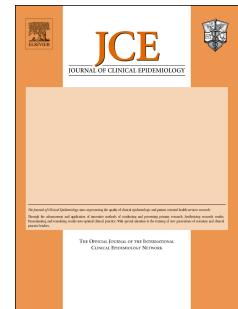
Received Date: 13 May 2016

Revised Date: 2 November 2017

Accepted Date: 9 November 2017

Please cite this article as: Cargo M, Harris J, Pantoja T, Booth A, Harden A, Hannes K, Thomas J, Flemming K, Garside R, Noyes J, Cochrane Qualitative and Implementation Methods Group Guidance Paper 3: Methods for Assessing Evidence on Intervention Implementation, *Journal of Clinical Epidemiology* (2018), doi: 10.1016/j.jclinepi.2017.11.028.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Title. Cochrane Qualitative and Implementation Methods Group Guidance Paper 3: Methods for Assessing Evidence on Intervention Implementation

Author names and affiliations

Assoc/Prof Margaret Cargo^{1,2} (Corresponding author), Dr Janet Harris³, Assoc/Prof Tomas Pantoja⁴, Dr Andrew Booth⁵, Prof Angela Harden⁶, Assoc/Prof Karin Hannes⁷, Prof James Thomas⁸, Dr Kate Flemming⁹, Dr Ruth Garside¹⁰, Prof Jane Noyes¹¹

¹Health Research Institute
University of Canberra
University Drive, 22-B17
Email: margaret.cargo@canberra.edu.au

²Centre for Population Health University of South Australia
8th Floor Office 310, SAHMRI Building (North Terrace)
Adelaide, Australia

³ School of Health and Related Research (SchHARR)
University of Sheffield
Regent Court, 30 Regent Street
Sheffield, UK
Email: janet.harris@sheffield.ac.uk

⁴ Department of Family Medicine, Faculty of Medicine
Pontificia Universidad Católica de Chile
Centro Médico San Joaquín, Av. Vicuna Mackenna 4686, Macul
Santiago, Chile
Email: tpantoja@med.puc.cl

⁵ School of Health and Related Research (SchHARR)
University of Sheffield
Regent Court, 30 Regent Street
Sheffield, UK
Email a.booth@sheffield.ac.uk

⁶ Institute for Health and Human Development
The University of East London
Stratford Campus
Water Lane
London, UK
Email: a.harden@uel.ac.uk

⁷ Social Research Methodology Group, Centre for Sociological Research
Faculty of Social Sciences
KU Leuven
Leuven, Belgium

Email: karin.hannes@kuleuven.be

⁸ UCL Institute of Education

University College London

20 Bedford Way

London, UK

Email: james.thomas@ucl.ac.uk

⁹ Department of Health Sciences,

Faculty of Science, University of York

Seebohm Rowntree Building

Heslington

York, UK

Email: kate.flemming@york.ac.uk

¹⁰ European Centre for Environment & Human Health

University of Exeter Medical School

Knowledge Spa, Royal Cornwall Hospital

Truro, Cornwall, UK

Email: r.garside@exeter.ac.uk

¹¹ School of Social Sciences

Bangor University

Bangor

Gwynedd, UK

Email: jane.noyes@bangor.ac.uk

Abstract

Objective: This article provides reviewers with guidance on methods for identifying and processing evidence to understand intervention implementation. Study Design and Setting: Strategies, tools and methods are applied to the systematic review process to illustrate how process and implementation can be addressed using quantitative, qualitative and other sources of evidence (i.e., descriptive textual, non-empirical). Results: Reviewers can take steps to navigate the heterogeneity and level of uncertainty present in the concepts, measures and methods used to assess implementation. Activities can be undertaken in advance of a Cochrane quantitative review to develop program theory and logic models that situate implementation in the causal chain. Four search strategies are offered to retrieve process and implementation evidence. Recommendations are made for addressing rigour or risk of bias in process evaluation or implementation evidence. Strategies are recommended for locating and extracting data from primary studies. The basic logic is presented to assist reviewers to make initial review level judgements about implementation failure and theory failure. Conclusion: Although strategies, tools and methods can assist reviewers to address process and implementation using quantitative, qualitative and other forms of evidence, few exemplar reviews exist. There is a need for further methodological development and trialling of proposed approaches.

Running Title: Methods for Assessing Evidence on Intervention Implementation

Keywords: Systematic reviews, process evaluation, implementation, Cochrane, qualitative evidence synthesis; mixed-method synthesis

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

103

Key findings:

Strategies, tools and methods are available to support reviewers to address process and implementation using qualitative and process evaluation evidence and other evidence from quantitative studies included in Cochrane reviews.

What this paper adds to what was known?

Cochrane quantitative reviews of interventions should include steps to identify, synthesise and then integrate evidence to address reach, dose, fidelity, co-intervention, contamination and the role of contextual factors on implementation.

What is the implication and what should change now?

Cochrane quantitative reviews use risk of bias tools to rule out evaluation failure. This guidance suggests that systematic reviewers use complementary tools to make informed judgements about implementation failure and theory failure to strengthen internal validity and enhance the uptake of review findings by decision-makers.

104

105

106 In 2013, the Cochrane Qualitative and Implementation Methods Group (CQIMG) expanded

107 its remit to include issues related to assessing implementation in systematic reviews of

108 interventions. The CQIMG focus on implementation complements the scope of work of the

109 Cochrane Effective Practice and Organisation of Care Group which undertakes systematic

110 reviews of educational, behavioural, financial, regulatory and organisational interventions

111 designed to improve health professional practice and the organisation of health care

112 services.

113 Implementation, conceptualized as a planned and deliberately initiated effort with the

114 intention to put an intervention into practice (1), occupies the space between the 'blueprint

115 for the intervention' (i.e., assumptions articulating how and why an intervention is supposed

116 to work) and the 'outcomes observed in practice'. Process evaluation investigates the

117 activities and internal dynamics of an intervention during its implementation to determine

118 how well an intervention operates (2, 3). This article provides reviewers with guidance on

119 how to approach process and implementation in a Cochrane quantitative review of the

effects of an intervention. Some of the issues discussed are relevant for both qualitative and quantitative reviews. This paper should be read in conjunction with the articles in this series about question formulation (4), evidence-appropriate methods for qualitative synthesis of evidence on implementation(5) and methods for integrating findings from qualitative syntheses with intervention effectiveness reviews (6), as it provides complementary information on how to refine implementation questions, retrieve process evaluation evaluations or implementation data and rule out implementation failure and theory failure when integrating the findings from qualitative syntheses with intervention effectiveness reviews.

Why is implementation important?

Too often quantitative reviews assess intervention outcomes (i.e., does it work) without considering how the process of implementation influences observed outcomes. In these reviews, causal inferences can be undermined from limitations in the design, data collection and analysis of primary studies and lead to an under- or overestimation of the true intervention effect. To assess the internal validity of primary quantitative studies, review authors apply risk of bias tools to make judgements about a number of methodological biases (i.e., selection, performance, detection, attrition, reporting) (7). Assessing risk of bias can rule out evaluation failure due to methodological biases that compromise internal validity (2). Although risk of bias is necessary to assess the strength of causal inferences in determining whether interventions are successful, it is not sufficient. Reviewers additionally need to establish the presence of a functional relationship between intervention implementation (i.e., independent variable) and a change in the outcome (i.e., dependent variable). To draw valid conclusions both need to be defined and evaluated. At a practical level information needs to be extracted from each primary study to inform a judgement about the integrity of implementation, and to examine whether specified procedures in the primary studies were implemented as outlined in the intervention protocols.

Formal evaluation of implementation in a process evaluation enables reviewers to determine whether key implementation outputs were achieved (8). Synthesising this information across primary studies can enhance the internal validity of systematic reviews by ruling out implementation failure and theory failure and provide decision-makers with insights into the conditions needed to generate positive outcomes in the target population

(8). Implementation failure is suspect when the lack of expected outcomes is attributed to poor implementation practices. Theory failure is suspect when intervention activities are implemented according to the specified standards, guidelines or intervention design strategy but expected outcomes are not observed. This suggests that the theory, logic or set of assumptions that specify how the intervention was expected to bring about change was incorrect (9). It is additionally important to consider the important role of contextual factors as interventions can be implemented and received differently in different contexts (10). Moreover, an unfavourable context can have a significant impact on the feasibility to implement or scale-up an intervention (11).

The example in Box 1 illustrates how the behavioural effects of a school-based program for children are influenced by implementation.

Box 1: Example highlighting the importance of accounting for implementation in quantitative reviews of interventions.

Aspects of implementation were accounted for in a systematic review that assessed the effects of universal school-based social information processing interventions on the aggressive and disruptive behaviour of school-age children(12). Studies reporting problems with program implementation produced smaller effect sizes compared to those not reporting such problems. Moreover, programs delivering more frequent treatment sessions per week were more effective than programs delivered less frequently. Review authors hypothesise that the cognitive skills emphasised by these types of programs may be hard to master and that more frequent delivery provides children with more opportunities for practice and reinforcement. These measures of implementation provide decision-makers with useful information on the conditions under which social information programs are more likely to reduce aggressive and disruptive behaviour in children.

What aspects of implementation are assessed and how?

Assessing implementation is a crucial component in the systematic reviews of quantitative health and social care interventions. Lack of information on intervention implementation

weakens internal validity and inhibits the translation and uptake of evidence by decision-makers to inform policy and practice. Aspects of implementation can be quantitatively assessed in different types of studies. These studies include randomised trials which answer questions pertaining to *“Can this intervention work in highly controlled or ideal conditions?”* positioned at the explanatory end of the pragmatic-explanatory spectrum (10) (i.e., ‘efficacy’ studies) and *“Does this intervention work in real world or usual care conditions?”* positioned at the pragmatic end of the pragmatic-explanatory spectrum (i.e., ‘effectiveness’ studies). Dissemination studies evaluate how the targeted distribution of intervention materials to a specific audience can be successfully implemented so the increased spread of knowledge about the evidence-based achieves greater use and impact of the evidence-based interventions(13). Implementation studies evaluate how a specific set of activities and designed strategies are used to successfully integrate and sustain an evidence-based interventions within specific settings (13). Scale-up studies evaluate deliberate efforts to increase the impact of evidence-based interventions to benefit more people and to foster policy and program development on as lasting basis(13). Policy analysis, which involves identifying the possible policy options to address a health and social care problem and then using the appropriate methods to determine the most effective, feasible and efficient option, is featured in dissemination, implementation and scale-up studies. In addition, it is increasingly common that qualitative ‘sibling’ studies and mixed-method process evaluations are undertaken alongside a trial, which can be synthesised to better understand the political and operational factors associated with the implementation of health policy, health systems, behavioural, environmental or clinical interventions. A synthesis of qualitative studies that are unrelated to trials can also be helpful in understanding the factors that affect intervention implementation (14, 15).

Process evaluations focus on one or more aspects of implementation, including reach, dose delivered, dose received, fidelity, adaptation, intervention quality, recruitment, provider engagement, participant engagement and contamination, co-intervention. Contamination and co-intervention are commonly included in risk of bias assessments (10, 16, 17). Table 1 provides definitions for these terms with example quantitative indicators and qualitative questions. At a minimum, it is recommended that a process evaluation includes information on reach, dose delivered/ received, fidelity and co-intervention, contamination (17) and

supplementary information on contextual factors (10, 17, 18). Including the latter in process evaluation aligns with the growing body of literature on complex interventions which recognises that intervention outcomes and implementation are highly influenced by contextual factors (1). The specific measures used to assess implementation in interventions will vary depending on whether reviews include efficacy, effectiveness, dissemination, implementation, policy or scale-up studies. The reason for this is that implementation is defined relative to the intervention content and as studies move from bench to bedside to population, the concepts of reach, dose and fidelity pertain to different aspects of the health and social care system. In complex reviews it is possible that these concepts may be assessed at two levels of the system (e.g., extent to which patients adhere to a treatment and the extent to which clinicians adhere to practice guidelines). In this regard, Harris (4) provides strategies for reviewers to apply in formulating review questions for complex interventions, which may include those with multiple implementation chains. We recommend review authors consider these dimensions as minimum requirements for inclusion in systematic reviews, and further consider reach, dose delivered/ received, fidelity and co-intervention, contamination as 'Other sources of bias' in the Cochrane 'Risk of bias' tool (7). When process evaluations in quantitative reviews are lacking, or results do not adequately address decision-makers concerns and qualitative perspectives on implementation are sought (Table 1) we recommend review authors collaborate with qualitative review teams to meet these minimum requirements (19).

Context-dependence of implementation

As a process, implementation is context-dependent and concerns the actions required to put an 'intervention blueprint' into practice (10). Context includes the immediate environment in which an intervention is implemented and broader environment that shapes the resources, political support and norms influencing engagement of the target audience (e.g., patients, practitioners). It can be difficult for reviewers to grasp these dimensions of implementation and locate them in a process evaluation. The UK Medical Research Council (MRC) Guidance on process evaluation of complex interventions provides a framework that links context, with the intervention description, implementation and the mechanisms of impact on outcomes (10). The framework in Figure 1 situates an intervention and its designated target populations in relation to the immediate and broader contexts within which the intervention is planned, implemented and evaluated (20). It can be used in conjunction with the MRC framework to help reviewers frame implementation in a formal logic model within their Cochrane review of quantitative interventions. The red line drawn around the intervention, target populations and program implementation boxes in Figure 1 visually depicts how resources and the external environment in addition to factors internal to the program environment (i.e., action model), are instrumental to shaping implementation. Box 2 illustrates how intervention outcomes can vary according to contextual factors.

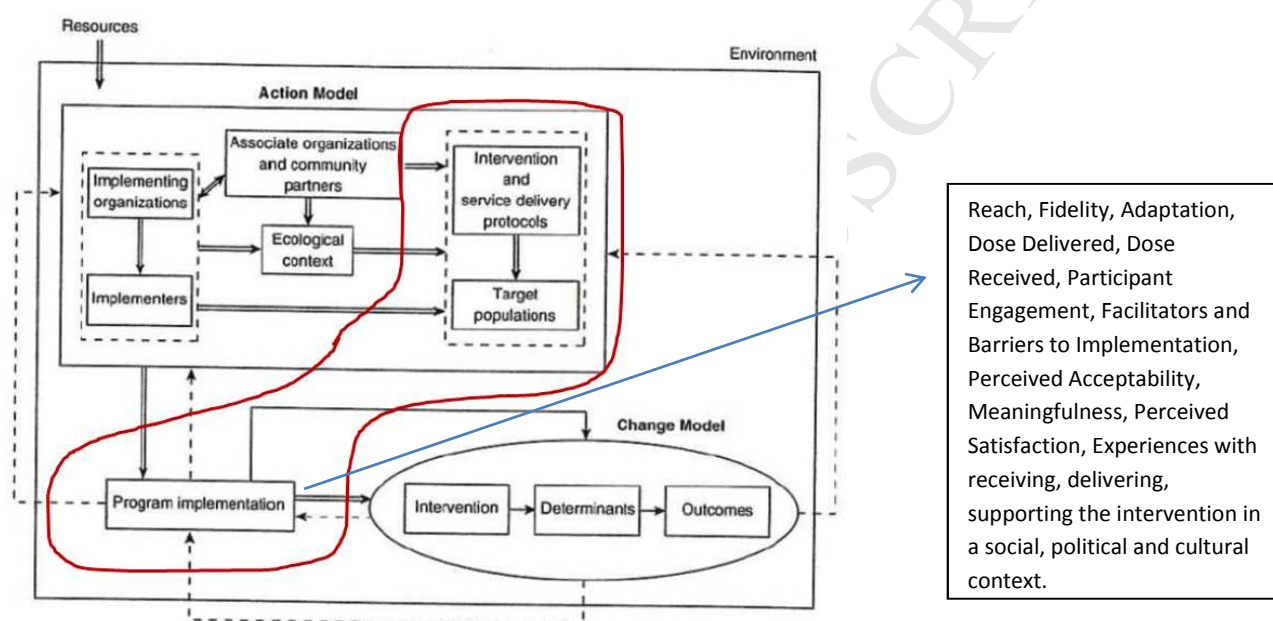
Box 2: Example of contextual factors influencing program outcomes

A meta-analysis of school-based programs to reduce bullying and victimisation found the impacts of these programs to vary by country of implementation (21). The programs worked better in Norway specifically and Europe more generally as compared to North America. The review authors posit that Scandinavian schools have a tradition of state intervention in social welfare and that the program context (i.e., high quality schools with small classes and well-trained teachers) may also contribute to the observed differences in outcomes.

Intervention delivery and service delivery protocols specify the nature, content and activities of an intervention, including its operating procedures, and the particular steps that

need to be taken to implement the intervention(20). This is the 'blueprint for the intervention'. What is implemented and how it is implemented to reach its designated target populations is documented through process evaluation. Implementation can be measured quantitatively through self-report surveys, structured observations, and secondary analysis of routine monitoring data or qualitatively through focus groups, individual interviews, unstructured observations (10) and open-ended survey questions.

Figure 1. Conceptual Framework to Situate Implementation in Relation to Context



Source: Chen H-T. Practical Program Evaluation. Thousand Oaks, CA: Sage Publications, 2005. Reprinted with permission from Sage Publications.

Steps of the Systematic Review Process

Increasingly, review authors of both quantitative and qualitative reviews are being called to address issues relevant to context and implementation to make the findings more applicable to decision-makers. We used the steps of the review process to illustrate how qualitative and other sources of evidence on implementation can be synthesised and then integrated with evidence of effect.

Step 1 - Framing the Problem and Refining Implementation Questions

The first step in a quantitative systematic review frames the problem and identifies which aspects of implementation are relevant. Framing the problem is driven by a number of

factors including the state of knowledge on a review topic, level of resourcing, timeframe, expertise, stakeholder input, and expectations from the review commissioners. Knowing where to start can be challenging for review authors especially if one or more of the following conditions is present: (a) there is considerable heterogeneity in the interventions considered for a review; (b) there is little understanding of how interventions work to produce outcomes for the population or context(s) of interest; (c) aspects of implementation are not clearly understood, are poorly defined or the evidence needed to address implementation cannot be clearly specified; (d) it is not clear how to frame the review question from an implementation perspective; or (e) stakeholders raise questions that are pertinent to implementation, and it is not clear how to address them. If one or more of these situations is apparent, we recommend a scoping review or other review activity with an implementation focus be undertaken, as outlined in Table 2, to help define or refine implementation issues and questions of interest (22) and inform a subsequent Cochrane systematic review of interventions. These methods align with current systematic review practices and guidance to formulate review questions that are inclusive of process and implementation issues (23, 24). Brief descriptions of the methods are provided in Appendix 1, available online as supplementary material (www.iclinepi.com).

Table 2: Strategies, methods and tools to help refine the questions and scope of a Cochrane effectiveness review.

Issue or circumstance	Review activity	Tools to assist	Product
When a broad range of interventions have been implemented to address a health issue.	Critical Review (25) Textual Narrative Synthesis(26)	Principles of simple, complicated and complex interventions (27); Template for Intervention	Classification of interventions; identification of program theory, logic model, implementation measures/ processes.
Lack of clarity in implementation concepts, definitions, measures or methods for a review.	Scoping Review [13] Concept Analysis(29)	Description and Replication (TiDIER) (28); Logic model template to situate implementation(23)	Implementation definitions for an effectiveness review; implementation concepts to assess in a qualitative synthesis.
An intervention model or framework for an effectiveness review requires adaptation to another topic or context.	Best-fit framework (30)	Logic model template to situate implementation(23)	Framework to guide the review with implementation situated in the framework.
Poor understanding of	Grounded Theory,	Logic model	Program theory and

program theory and how implementation relates to outcomes; review resources are available.	Realist Synthesis, Meta-Ethnography, Meta-Interpretation (24)	template to situate implementation (23)	logic model with implementation concepts and indicators identified.
As above, but review resources are not available.	Program theory mapping workshop	Logic model template; 'how-to' resources (27); engage consultant.	Program theory and logic model with implementation concepts identified.

Following Harris (4) and the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)(31, 32), we recommend that reviewers engage stakeholders in the preparatory stage to ensure that the review scope is appropriate and resulting products address the implementation inquiry questions and concerns of decision-makers. These review activities will increase the internal validity of constructs, measures and methods used in a quantitative review which can reduce the likelihood of evaluation failure and strengthen the basis for making judgements that rule out implementation and theory failure.

Step 2 - Searching

As shown in Table 1, a search for the following types of evidence may potentially help with understanding intervention implementation:

- 'Implementation evidence' from quantitative studies (e.g. RCTs included in the effect review) on dose and reach etc.
- 'Process evaluation evidence' - qualitative and quantitative evidence from process evaluations conducted alongside trials
- 'Trial sibling qualitative studies' – conducted alongside trials
- 'Unrelated qualitative studies' – with no relationship to trials
- Economic evaluations – conducted alongside trials

Retrieval of process evaluations and implementation evidence of all types is problematic for at least three reasons. First, process evaluations may not exist. Second, when they do exist, they may not be clearly identifiable in terms of key terms for their retrieval. Third, process evaluations may not be published in the peer reviewed literature (33) and, therefore, carry the challenges associated with retrieving grey or fugitive literature (34). The CQIMG has identified four potential approaches to identify process evaluations in a systematic review (35). The approach that is used will be determined by factors such as the review purpose, time and resource constraints and the perceived risk of how deficiencies in the search

process will impact upon the uncertainty of the review results. The first approach for retrieving process evaluations is to transfer identification from the search process to the sift process. This involves conducting a sensitive topic search without any publication restrictions (36). The review team works its way systematically through the titles and abstracts of retrieved references looking for indications of process data by using the dimensions highlighted in Table 1. This approach is feasible when a review question involves multiple publication types e.g. RCT, qualitative research and economic evaluations, which are not being searched for separately. The second approach retrieves process evaluations within randomised control trials for which the Cochrane has developed a highly sensitive search strategy (filter) (37). If a process evaluation has been published in a journal article and mentions the trial in the abstract, this method proves effective. The third option is to use unevaluated filter terms to retrieve process evaluations or implementation data. Approaches using strings of terms associated with the study type or purpose is considered experimental. There is a need to develop and test such filters. It is likely that such filters may be derived from the study type (process evaluation), the data type (process data) or the application (implementation). The last of these is likely to prove problematic because a study can describe implementation without necessarily using the word “implementation” (38). The fourth approach relies on citations-based approaches. We have proposed the identification of ‘clusters’ containing all accounts, published or unpublished, of a particular study (39). These can offer additional contextual detail but, importantly in this context, may provide implementation or process data (40).

At present, the CQIMG suggests that review teams either use methods 2 and 3 in conjunction with 4, most likely in a Cochrane setting, or use method 1 in conjunction with 4 for a wider health technology assessment type ‘multi-review’ (35). Guidance on searching for trials can be found in the Cochrane Handbook (37) and paper 2 in this series outlines principles for searching for qualitative studies (5).

Step 3 - Data Extraction

To extract relevant information on implementation from primary studies it is crucial to have a detailed understanding of the intervention because implementation measures (e.g., fidelity, dose) and the barriers and facilitators experienced during implementation can

pertain to different aspects of complex interventions (10, 17). We therefore recommend use of the 10-dimension Complexity Assessment Tool for Systematic Reviews (iCAT-SR) to assist with classifying and grouping interventions (41). For quantitative intervention reviews, this can inform sub-group or sensitivity analyses, and aid in developing logic models and identifying causal pathways that explicitly feature implementation (Lewin, forthcoming). For qualitative evidence syntheses, the iCAT-SR may facilitate comparisons of staff experiences with implementation or the construction of implementation chains for different types of programs, enhancing the theoretical and interpretive validity of the review.

A review of 27 systematic reviews of interventions uncovered several issues impacting the extraction of information on implementation from primary studies (42). Process evaluation terms are not always defined and reviewers may find aspects of implementation described (i.e., ‘the evaluation assessed whether the intervention was implemented as intended’) but not linked to a specific definition (i.e., fidelity). Terms or definitions are not located in the methods section which is where review authors might expect to find them; sometimes they appear in the discussion section. Aspects of implementation are defined in ways that deviate from commonly accepted definitions. For example, studies can define intervention ‘quality’ as the intervention being delivered as intended, which is the definition commonly used for fidelity (43). Like the intervention, information on program operations (‘implementation’) is often descriptive (i.e, textual) and not empirical and can appear in the background and methods section of a primary outcome evaluation paper, or in a non-empirical ‘sibling’ study. Additionally, authors often provide reflections on implementation in the discussion section. To counteract some of these limitations, following the techniques used in Intervention Component Analysis (44) we recommend that descriptive information and author reflections on the experience of implementing the intervention are used from trial and ‘sibling’ reports and further, that corresponding authors be contacted for specific information on implementation. Such information strengthens the descriptive validity of qualitative and quantitative reviews. We also recommend that review authors develop a glossary of terms and definitions supported by existing resources such as the Oxford Implementation Index (45), Checklist for Implementation (42) and the MRC Guidance on process evaluation of complex intervention (10) to reduce the likelihood of conceptual

slippage and inconsistent interpretation of measures of events between studies. For systematic reviews, this can guide the consistent extraction of information across studies. For a qualitative evidence synthesis, a common set of understandings of key implementation terms and processes can facilitate comparisons of experiences between studies which, again, can enhance theoretical and interpretive validity.

Step 4 – Assessing Rigour and Risk of Bias in the process evaluation or intervention implementation evidence

Review authors should determine if the absence of a favourable intervention effect within primary studies and at the review level is due to problems with implementation (i.e., implementation failure) or a poorly conceptualised intervention (i.e., theory failure). Few assessment tools for primary studies or reviews explicitly address the rigour or risk of bias in process evaluation or implementation evidence. Table 1 in Noyes et al (this series (5)) reports comparable terms (such as risk of bias and rigour) to describe similar domains across quantitative and qualitative research. Building on previous recommendations(46), we provide recommendations for assessing the rigour/risk of bias of process and implementation in primary studies and reviews.

The literature was systematically searched to retrieve tools to critically appraise process and implementation. This entailed keyword searches of PubMed MEDLINE, the ISI Web of Science, the worldwide web, Google Scholar, the webpages of systematic review centres/ collaborations and perusing the reference lists of relevant documents. This search was initially conducted in 2009(47) and updated periodically through CQIMG-affiliated work. One assessment tool specific to process evaluation was located. This 8-item tool developed by the EPPI-Centre is flexible and can be applied to qualitative, quantitative and mixed-method primary studies (48, 49). Six questions tap rigour related to sampling, data collection, data analysis, interpretation, breadth/scope of findings, and whether the study privileges the perspective of the target group. The last two items assess the reliability and usefulness of the findings. The question on usefulness (*'how well the intervention processes were described and whether or not the process data could illuminate why or how the interventions worked or did not work'*) offers insight into process mechanisms. Ideally process evaluation should gather both qualitative and quantitative information. Qualitative

data is particularly important to understand how features of context influence implementation and issues related to acceptability, meaningfulness and generalisability of the intervention. As outlined below, we recommend this 8-item tool supplement existing critical appraisal tools for primary qualitative and quantitative studies. Given that existing critical appraisal tools for systematic reviews do not address process evaluation and following recent guidance on the process evaluation of complex interventions(10) we recommend that questions be developed to supplement these tools.

For qualitative primary studies we recommend the 8-item process evaluation tool (49) be used in conjunction with a qualitative critical appraisal tool such as the Evaluation Tool for Qualitative Studies (ETQS)(50). The ETQS was the only tool of three qualitative tools reviewed to cover all forms of validity (i.e., descriptive, theoretical, evaluative, interpretive, generalisability)(51) and it additionally enquires into study context, specifically setting factors and the sampling of events, persons, times and settings both of which are important to understanding implementation. While the process evaluation specific tool captures rigour relevant to implementation, the ETQS captures rigour relevant to qualitative validity (credibility and transferability). These tools should be used in addition to tools to assess methodological strengths and limitations that feed into CERQual assessments of confidence in synthesised qualitative findings (52).

Assessment tools for quantitative primary studies do not address dimensions of process evaluation other than contamination, co-intervention, and participation. The Effective Public Health Practice Project Quality Assessment Tool (EPHPP) (53) is the only tool that asks a question on fidelity, operationalised as consistency of implementation. Overall integrity is judged by responses to three questions on fidelity, contamination/ co-intervention and percentage of participants receiving the allocated intervention. The Cochrane Risk of Bias Tool (CRBT) was introduced to establish consistency and avoid discrepancies in the assessment of methodological strengths and limitations. Considering that Cochrane reviewers are required to use the CRBT we recommend its use be supplemented with the 8-item process evaluation assessment tool (49). This tool is flexible and allows Cochrane reviewers to make an assessment of the methodological strengths and limitations of an

embedded or sibling process evaluation study that includes one or more of the dimensions in Table 1 using quantitative, qualitative or mixed methods

Step 5 – Analysis, Synthesis and Interpreting the Evidence with an Implementation Lens

Papers 2(5) and 4(6) in the series provide an overview of evidence-appropriate methods for synthesis of evidence on implementation, and paper 4 outlines methods for integrating qualitative and process evaluation evidence with evidence of intervention effect.

At the final stage, evidence from the qualitative and quantitative reviews need to be brought together to inform a judgement about ‘implementation success or failure’ and ‘theory success or failure’ (either partial or complete) at the integrated review level. At present no Cochrane reviews of interventions formally do this, however, information, in some reviews allows for less formal retrospective or ad-hoc judgements of theory failure and implementation failure (Box 3).

Box 3: Ruling out implementation failure and theory failure

Petrosino et al (54) reviewed the effects of programs comprised of organised visits to prisons by juvenile delinquents or pre-delinquents to deter them from delinquency (‘Scared Straight’). The meta-analysis found the organised prison visits to be more harmful than doing nothing. Problems with implementation were considered as a potential source of bias. All included studies were considered low risk of bias as no investigator reported problems with implementation. Since the programs were implemented with fidelity, the harmful effect suggests fault in the program’s logic that exposing at-risk juveniles to prison life would deter delinquency. The authors posit peer contagion theory as a potential explanation for the observed effect; the potential intervention benefit was offset by deviant youth interacting with each other in a group setting. This alternate causal pathway could be explored in a qualitative evidence synthesis.

We argue that reviews need to be designed at the problem description stage to address this, specifically by generating a program theory or logic model that depicts implementation outputs or measures captured quantitatively, or core processes captured qualitatively. The

basic logic for informing such judgements is outlined in Figure 2a-c. Implementation failure and theory failure do not operate in isolation. To determine whether theory failure is suspect in interpreting the overall intervention effect of a primary study, it is necessary to first rule out implementation failure. If a review does not systematically extract qualitative and/or quantitative evidence on implementation and finds that the primary outcome did not favour the treatment condition, reviewers do not have a basis for determining, at the interpretation stage, whether the intervention design was deficient (theory failure) or whether the outcome was marred due to implementation problems (implementation failure). This compromises the overall internal and external validity of the review. The example in Box 3 additionally highlights the need to assess implementation in order to be able to make a judgement about underlying program theory.

The activities in Table 2 increase the chance that reviews are guided by plausible and testable program theory. The MRC Process Evaluation Framework(10) and the framework outlined in Figure 1 provides reviewers with the conceptual building blocks to develop program theory. For any given review, program theory visually depicted in a logic model acts as a 'coat rack' of sorts to hang the most appropriate measures and methods to capture the uniqueness of intervention contexts in primary studies. Hence, context becomes 'reproducible' by virtue of the conceptual frameworks, methods, measures and tools used to construct the logic that guide reviews. The synthesis methods described in papers 2(5) and 4(6) in the series provide insight into differential intervention effects, context by implementation interactions and inform judgements about partial or complete breakdowns in implementation. Methodological work is required to inform review level judgements of implementation and theory failure, whether partial or complete.

467 Figure 2a-c¹. Depictions of implementation and theory failure

Implementation failure is ‘diagnosed’ by determining whether intervention activities produce the requisite operation outputs, depicted as the first intervening variable in Figure 2a. These outputs pertain to key implementation measures (e.g., dose delivered, reach, fidelity) and processes. If these outputs are not achieved the causal pathway has been disrupted and we wouldn’t expect to see a change in the short-term goal or bridging variable, or the primary outcome.

Figure 2a. Implementation Failure



Theory failure is suspect when a process evaluation shows that an intervention achieved its key operation outputs (i.e., intervention implemented with integrity) but not its short-term goal (e.g., increase in physical activity), depicted as the intervening bridging variable in Figure 2b.

Figure 2b. Theory Failure (Case 1)



Theory failure is also suspect when an intervention achieves its operation outputs (i.e., implementation integrity) and short-term goal (e.g., increase in physical activity) but the short-term goal or bridging variable doesn’t translate to a change in the primary outcome (e.g., body mass index) (Figure 2c).

Figure 2c. Theory Failure (Case 2)



468 ¹Adapted from (9)

Conclusions

Assessing implementation in Cochrane systematic reviews of interventions is challenging for a number of reasons, including, but not limited to, poor reporting of intervention and implementation in primary studies, knowing the starting point to address implementation on a given topic, and pressures to accommodate knowledge translation concerns of research consumers despite reporting and review resource limitations. Depending on the review objectives, synthesis of evidence on implementation can add interpretive value to Cochrane reviews and the decision-makers who use them. This paper provides guidance for reviewers to navigate the heterogeneity and uncertainty that they are confronted with at different stages of the review process.

Table 1: Definitions of key dimensions of implementation with corresponding examples of quantitative indicators and qualitative questions.

Dimension	Quantitative	Qualitative
Dose Delivered: Amount of a program delivered to participants (i.e., frequency, duration, intensity) by staff and/or implementing agency.	<ul style="list-style-type: none"> • Total # contact hours • # water fountains installed 	<ul style="list-style-type: none"> • How did participants feel about the format and time commitment of the program?
Dose Received: Characteristic of the target population's utilisation or interaction with program strategies or resources ('active participation').	<ul style="list-style-type: none"> • Dosage of medicine ingested • # people drinking water from fountain 	<ul style="list-style-type: none"> • What factors influenced whether clients read the take home educational materials?
Reach: Degree to which target group participates by their presence.	<ul style="list-style-type: none"> • # of patients served by eligible clinics 	<ul style="list-style-type: none"> • What motivated clients to attend the clinic?
Recruitment: Specific information on procedures used to recruit or attract participants to the intervention.	<ul style="list-style-type: none"> • % of clients recruited by type of recruitment strategy 	<ul style="list-style-type: none"> • How did participants feel about the methods used to recruit them?
Fidelity: Reflects implementation integrity, adherence, extent to which a program is implemented as intended.	<ul style="list-style-type: none"> • % of activities critical to behaviour change completed 	<ul style="list-style-type: none"> • What factors enabled clinical staff to adhere to practice guidelines?
Adaptation: Whether aspects of a program were intentionally changed during delivery to enhance outcomes.	<ul style="list-style-type: none"> • % of activities that changed during intervention period 	<ul style="list-style-type: none"> • What factors influenced staff adaptation of intervention activities?
Co-intervention: When interventions other than the treatment are applied differently to intervention conditions.	<ul style="list-style-type: none"> • % of control group participants getting other treatments 	<ul style="list-style-type: none"> • Why did participants engage in other activities related to the outcome?
Contamination: Unintentional delivery of intervention to the control group or inadvertent failure to deliver intervention to experimental group.	<ul style="list-style-type: none"> • % of control group participants exposed to the treatment 	<ul style="list-style-type: none"> • How did the control group come to receive the treatment?
Participant Engagement: Participant's interaction with or receptivity to a program i.e., what they think or how they feel about the intervention	<ul style="list-style-type: none"> • On a scale of 1 to 5, rate the extent to which the program met your needs 	<ul style="list-style-type: none"> • Was the program culturally appropriate and acceptable to clients?
Implementer Engagement: Subjective	<ul style="list-style-type: none"> • On a scale of 1 to 5, 	<ul style="list-style-type: none"> • How would you

staff attributes that influence program delivery i.e., what they think/ feel about the intervention and their interpersonal style.	rate your level of enthusiasm to use the practice guidelines	characterise your motivations and interests to implement the practice guidelines?
Intervention Quality: Quality of intervention materials/ resources (e.g., curriculum, training, policy).	<ul style="list-style-type: none"> On a scale of 1 – 5 rate the quality of the training 	<ul style="list-style-type: none"> Please comment on the training materials and facilitation of the training
Context: Social, built and political factors internal (e.g., partnerships) and external to the intervention environment (e.g., social norms) that shape implementation.	<ul style="list-style-type: none"> On a scale of 1 – 5, to what extent did community agencies support the intervention? 	<ul style="list-style-type: none"> In what ways did community agencies support the health service to deliver the intervention?

1. Pfadenhauer LM, Mozygemba K, Gerhardus A, Hofmann B, Booth A, Lysdahl KB, et al. Context and implementation: A concept analysis towards conceptual maturity. *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen*. 2015;109(2):103-14.
2. Rossi PH, Lipsey, M.W., Freeman, H.E. *Evaluation. A Systematic Approach*. Seventh ed. Thousand Oaks: Sage Publications; 2004.
3. Patton MQ. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. Los Angeles: Sage; 2015. 805 p.
4. Harris J, Booth, A, Cargo, M, Hannes K, Harden A, Flemming K, Garside R, Pantoja T, Thomas J, Noyes J. Cochrane Qualitative and Implementation Methods Group Guidance Paper 1: Methods for question formulation, identifying and processing evidence on intervention. *Journal of Clinical Epidemiology*. forthcoming.
5. Noyes J, Booth A, Flemming K., Garside R, Lewin S, Harden A, Pantoja T, Hannes K, Cargo M, Thomas J. Cochrane Qualitative and Implementation Methods Group Guidance Paper 2: Methods for assessing methodological limitations, data extraction and synthesis, and confidence in synthesized qualitative findings. *Journal of Clinical Epidemiology*. forthcoming.
6. Harden A, Thomas, J, Cargo, M, Harris J, Pantoja, T, Flemming K, Booth A, Garside R, Hannes K, Noyes J. Cochrane Qualitative and Implementation Methods Group Guidance Paper 4: Methods for integrating findings from syntheses of qualitative and process evaluation evidence within intervention effectiveness reviews. *Journal of Clinical Epidemiology*. forthcoming.
7. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
8. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*. 1998 18(1):23-45.
9. Weiss CH. *Evaluation Research. Methods of Assessing Program Effectiveness*. Smelser HCN, editor. Englewood Cliffs, New Jersey: Prentice-Hall; 1972.
10. Moore G, Audrey S, Barker M, Lyndal B, Bonell C, Hardeman W, Moore L, O'Cathain A, Tinati T, Wight D, Baird J. *Process Evaluation of Complex Interventions*. UK Medical Research Council (MRC) Guidance.; 2015.
11. Milat A, Newson R, King L, Rissel C, Wolfenden L, Bauman A, Redman S, Giffin M. A guide to scaling up population health interventions. *Public Health Research & Practice*. 2016;26(1).
12. Wilson SJ, Lipsey MW. The effects of school-based social information processing interventions on aggressive behavior: Part I: Universal programs. *Campbell Systematic Reviews* 2006.
13. Schillinger D. An introduction to effectiveness, dissemination and implementation research. A resource manual for community-engaged research. University of California San Francisco Clinical and Translational Science Institute (CTSI) Community Engagement Program; 2010.
14. Noyes J, Hendry M, Lewin S, Glenton C, Chandler J, Rashidian A. Qualitative "trial-sibling" studies and "unrelated" qualitative studies contributed to complex intervention reviews. *J Clin Epidemiol*. 2016.
15. Moore G, Audrey S, Barker M, Bond L, Bonell C, Cooper C, et al. Process evaluation in complex public health intervention studies: the need for guidance. *Journal of Epidemiology and Community Health*. 2013.
16. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation Science*. 2007;2:40.
17. Steckler A, Linnan, L. (Eds) *Process Evaluation for Public Health Interventions and Research*. San Francisco: Jossey-Bass; 2002.
18. Waters E, de Silva-Sanigorski A, Hall B, Brow T, Campbell K, Gao Y, Armstrong R, Prosser L, & Summerbell C. Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews*. 2011;12(No.: CD001871, p. 10.1002/14651858.CD14001871.pub14651853.).
19. Creswell JW, Klassen AC, Plano Clark VL, Smith KC. for The Office of Behavioral and Social Sciences Research. *Best Practices for Mixed Methods Research in the Health Sciences*. National Institutes of Health; 2011.

20. Chen H-T. Practical Program Evaluation. Assessing and Improving Planning, Implementation and Effectiveness. Thousand Oaks, CA: Sage Publications; 2005.
21. Farrington D, Ttofi M. School-Based Programs to Reduce Bullying and Victimization. Campbell Systematic Reviews 2009.
22. Noyes J, Pearson A, Hannes K, Booth A on behalf of the Cochrane Qualitative Research Methods Group. Qualitative Research and Cochrane Reviews. In: J. Higgins & S Green (Eds). Cochrane Handbook for the Systematic Review of Interventions: Cochrane Book Series. UK: John Wiley & Sons Ltd; 2008. p. 571-92.
23. Anderson LM, Petticrew M, Rehfuss E, Armstrong R, Ueffing E, Baker P, et al. Using logic models to capture complexity in systematic reviews. Research Synthesis Methods. 2011;2(1):33-42.
24. Booth A, Noyes J, Flemming K, Gerhardus A, Wahlster P, Van der Wilt GJ, Mozygemba K, Refolo P, Sacchini D, Tummers M, Rehfuss E. Guidance on choosing qualitative evidence synthesis methods for use in health technology assessments of complex interventions [Online]. Available from: <http://www.integrate-hta.eu/downloads/> 2016.
25. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health information and libraries journal. 2009;26(2):91-108.
26. Barnett-Page E, Thomas J. Methods for the synthesis of qualitative research: a critical review. BMC Med Res Methodol. 2009;9:59.
27. Funnell SC, and Rogers, PJ. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models: John Wiley/Jossey-Bass; 2011.
28. Hoffmann T, Glasziou P, Boutron I, Milne R, Perera R, Moher D. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348.
29. Walker LO, Avant, KC. Strategies for theory construction in nursing. Upper Saddle River, NJ: Pearson Prentice Hall; 2005.
30. Carroll C, Booth A, Leaviss J, Rick J. "Best fit" framework synthesis: refining the method. BMC Med Res Methodol. 2013;13:37.
31. Oliver K, Rees R, Brady LM, Kavanagh J, Oliver S, Thomas J Broadening public participation in systematic reviews: a case example involving young people in two configurative reviews. Research Synthesis Methods 2015;6(2):206-17.
32. Jamal F, Langford R, Daniels P, Thomas J, Harden A, Bonell C. Consulting with young people to inform systematic reviews: an example from a review on the effects of schools on health. Health Expectations: an international journal of public participation in health care and health policy. 2014.
33. Lewin S, Glenton C, Oxman AD. Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. BMJ. 2009;339:b3496.
34. Adams RJ, Huff AS. Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies. International Journal of Management Reviews. 2016; April 1.
35. Booth A. Identifying evidence on process and implementation Unpublished report. 2016.
36. Harden A, Oakley, A, & Weston, R. A review of the effectiveness and appropriateness of peer-delivered health promotion for young people.: Institute of Education, University of London; 1999.
37. Lefebvre CME, Glanville J. Chapter 6: Searching for studies. In: J. Higgins & S Green (Eds). Cochrane Handbook for Systematic Reviews of Interventions Version 5.10 (updated March 2011) The Cochrane Collaboration. Available from www.cochrane-handbook.org 2011.
38. Cooper Robbins SC, Ward K, Skinner SR. School-based vaccination: a systematic review of process evaluations. Vaccine. 2011;29(52):9588-99.
39. Booth A, Harris J, Croot E, Springett J, Campbell F, Wilkins E. Towards a methodology for cluster searching to provide conceptual and contextual "richness" for systematic reviews of complex interventions: case study (CLUSTER). BMC Medical Research Methodology. 2013;13:118.
40. Bonell C, Farah J, Harden A, Wells H, Parry W, Fletcher A, Murphy S. Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. Public Health Research. 2013;1(1).
41. Lewin S, Noyes J, Chandler J, Hendry M et al. Guidance for using the iCAT-SR: Intervention Complexity Assessment Tool for Systematic Reviews, version 1. Cochrane Collaboration. forthcoming.

42. Cargo M, Stankov I, Thomas J, Saini M, Rogers P, Mayo-Wilson E, et al. Development, inter-rater reliability and feasibility of a checklist to assess implementation (Ch-IMP) in systematic reviews: the case of provider-based prevention and treatment programs targeting children and youth. *BMC Medical Research Methodology* 2015;15(1):73.
43. Salmon J, Ball K, Crawford D, Booth M, Telford A, Hume C, et al. Reducing sedentary behaviour and increasing physical activity among 10-year-old children: overview and process evaluation of the 'Switch-Play' intervention. *Health Promotion International* 2005;20(1):7-17.
44. Sutcliffe K, Thomas J, Stokes G, Hinds K, Bangpan M. Intervention Component Analysis (ICA): a pragmatic approach for identifying the critical features of complex interventions. *Systematic Reviews*. 2015;4(1):1-13.
45. Montgomery P, Underhill K, Gardner F, Operario D, Mayo-Wilson E. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology* 2013;66(8):874-82.
46. Armstrong R WE, Jackson N, Oliver S, Popay J, Shepherd J, Petticrew M, Anderson L, Bailie RBG, Hawe P, Kristjansson E, Naccarel la L, Norris S, Pienaar E, Roberts H, Rogers W SA, Thomas H. Guidelines for Systematic reviews of health promotion and public health interventions. Version 2. Melbourne University: Australia; 2007
47. Cargo M. Tools for critically appraising process and implementation in systematic reviews. Paper presented at The Joanna Briggs Institute Biennial International Convention Ripples to Revolution: From Bench to Bedside; Adelaide, Australia.2009.
48. Rees R, Oliver K, Woodman J, Thomas J. Children's views about obesity, body size, shape and weight. A systematic review.: Social Science Research Unit, Institute of Education, University of London; 2009.
49. Shepherd J, Kavanagh J, Picot J, Cooper K, Harden A, Barnett-Page E, et al. The effectiveness and cost-effectiveness of behavioural interventions for the prevention of sexually transmitted infections in young people aged 13-19: a systematic review and economic evaluation. *Health Technology Assessment*. 2010;14(7):1-206, iii-iv.
50. Health Care Practice Research and Development Unit. Evaluation Tool for Qualitative Research. Retrieved from http://usirsalfordacuk/12970/1/Evaluation_Tool_for_Qualitative_Studiespdf. 2009.
51. Hannes K, Lockwood C, Pearson A. A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. *Qual Health Res*. 2010;20(12):1736-43.
52. Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gulmezoglu M, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med*. 2015;12(10):e1001895.
53. Effective Public Health Practice Project Quality Assessment Tool. Quality Assessment Tool for Quantitative Studies. 2009.
54. Petrosino A, Turpin-Petrosino C, Hollis-Peel ME, Lavenberg JG. 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews*. 2013(4).